Abstract:

Automating the search for optimal neural network architectures has become increasingly crucial with the proliferation of deep learning applications. This paper explores the fusion of Neural Architecture Search (NAS) and Reinforcement Learning (RL) to address the inefficiencies of traditional NAS methods. The proposed approach aims to streamline the architecture discovery process, ensuring a better balance between model accuracy, computational resources, and interpretability.

The literature review section delves into recent advancements in NAS using RL, emphasizing the state-of-the-art works that have significantly contributed to the field.

Literature Review:

In recent years, Neural Architecture Search (NAS) using Reinforcement Learning (RL) has witnessed remarkable progress, revolutionizing the design of neural networks. Below, we highlight ten influential papers that have contributed to the state-of-the-art in this domain:

1. "Neural Architecture Search with Reinforcement Learning" (Zoph and Le, 2017)

- This pioneering work introduced the concept of using RL to search for neural architectures, setting the stage for subsequent developments.

2. "Progressive Neural Architecture Search" (Liu et al., 2018)

- This paper proposed a method that progressively refines architectures, significantly reducing the computational cost of NAS.

3. "Efficient Neural Architecture Search via Parameter Sharing" (Pham et al., 2018)

- It introduced the idea of sharing parameters among candidate architectures, further enhancing the efficiency of NAS.

4. "DARTS: Differentiable Architecture Search" (Liu et al., 2019)

- DARTS pioneered differentiable NAS, enabling the direct optimization of architectural parameters.

5. "ProxylessNAS: Direct Neural Architecture Search on Target Task and Hardware" (Cai et al., 2019)

- This paper addressed the gap between search and deployment by considering hardware constraints during NAS.

6. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks" (Tan and Le, 2019)

- EfficientNet proposed a novel scaling method that achieved state-of-the-art performance with fewer resources.

7. "Once for All: Train One Network and Specialize it for Efficient Deployment" (Cai et al., 2020)

- This work introduced a single model that can be specialized for various tasks, improving efficiency.

8. "AutoML-Zero: Evolving Machine Learning Algorithms from Scratch" (Real et al., 2020)

- AutoML-Zero extended NAS concepts to evolve machine learning algorithms themselves, demonstrating NAS's versatility.

9. "Faster Neural Architecture Search with Reinforcement Learning" (Xie et al., 2020)

- This paper presented an accelerated NAS method, reducing the search time while maintaining performance.

10. "Once-for-All: Efficient Constraint-Aware Training of Once-for-All Networks" (Huang et al., 2021)

- Building upon the "Once for All" concept, this work introduced constraint-aware training, optimizing models for specific tasks and resources.

These papers collectively illustrate the rapid evolution of NAS using RL, showcasing its potential to revolutionize deep learning model design and deployment, making it more efficient, accessible, and adaptable to a wide range of applications.